DOCUMENT RESUME

ED 451 643 EC 308 325

AUTHOR	Kosciolek, Stacey; Ysseldyke, James E.
TITLE	Effects of a Reading Accommodation on the Validity of a Reading Test. Technical Report 28.
INSTITUTION	Council of Chief State School Officers, Washington, DC.; National Association of State Directors of Special Education, Alexandria, VA.; National Center on Educational Outcomes, Minneapolis, MN.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE	2000-12-00
NOTE	27p.
CONTRACT	H326G000001
AVAILABLE FROM	National Center on Educational Outcomes, University of Minnesota, 350 Elliott Hall, 75 East River Rd., Minneapolis, MN 55455 ($10). Tel: 612-624-8561; Fax: 612-624-0879; Web site: http://www.coled.umn.edu/NCEO.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Accommodations (Disabilities); Auditory Stimuli; *Disabilities; Elementary Education; *Performance Factors; *Reading Aloud to Others; *Reading Comprehension; *Reading Tests; Student Attitudes
IDENTIFIERS	*Testing Accommodations (Disabilities)

ABSTRACT

This study examined the effects of a reading accommodation on the performance of students on a reading comprehension test. In total, 17 general education students and 15 special education students in grades 3-5 participated. Each student took two equivalent forms of the California Achievement Tests (CAT/5), Comprehension Survey. One form was administered with a read aloud accommodation, the other was administered without it. The read aloud accommodation was provided with an audiocassette player. One male and one female adult reader alternated between each passage and set of items for each test form. Each passage was read once at a speed of approximately 120 words per minute. A signification interaction between accommodation effects by student status was not found. However, there was a moderate positive effect size for students with disabilities, while the effect size for students in the general population was minimal. In addition to the test scores, data on student preferences were collected. Students in special education preferred the test with the accommodation, students in general education preferred the test without the accommodation. The study concludes that further investigation with larger groups is needed to determine whether the trends found are maintained. (Contains 26 references.) (CR)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Technical Report 28

ED 451 643

Effects of a Reading Accommodation on the Validity of a Reading Test

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EC 308325

NATIONAL CENTER ON EDUCATIONAL OUTCOMES

In collaboration with:

Council of Chief State School Officers (CCSSO)

National Association of State Directors of Special Education (NASDSE)

BEST COPY AVAILABLE

Technical Report 28

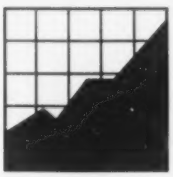
Effects of a Reading Accommodation on the Validity of a Reading Test

Stacey Kosciolek • James E. Ysseldyke

December 2000

All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Kosciolek, S., & Ysseldyke, J. (2000). *Effects of a reading accommodation on the validity of a reading test* (Technical Report 28). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.



NATIONAL CENTER ON EDUCATIONAL OUTCOMES

The Center is supported through a Cooperative Agreement (#H326G000001) with the Research to Practice Division, Office of Special Education Programs, U.S. Department of Education. The Center is affiliated with the Institute on Community Integration at the College of Education and Human Development, University of Minnesota. Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

NCEO Core Staff

John S. Bielinski
Jane L. Krentz
Michael L. Moore
Rachel F. Quenemoen
Dorene L. Scott
Sandra J. Thompson
James E. Ysseldyke

Martha L. Thurlow, Director

Additional copies of this document may be ordered for $10.00 from:

National Center on Educational Outcomes
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
Phone 612/624-8561 • Fax 612/624-0879
http://www.coled.umn.edu/NCEO

The University of Minnesota is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

This document is available in alternative formats upon request.

Executive Summary

The purpose of this study was to examine the effects of a reading accommodation on the performance of students on a reading comprehension test. In total, 17 general education students and 15 special education students, in grades 3-5 participated. Each student took two equivalent forms of the California Achievement Tests (CAT/5), Comprehension Survey. One form was administered with the accommodation; the other was administered without it. A significant interaction between accommodation effects by student status was not found. However, there was a moderate positive effect size for students with disabilities, while the effect size for students in the general population was minimal. In addition to the test scores, data on student preference were collected. Students in special education preferred the test with the accommodation, students in general education preferred the test without. Further investigation with larger groups is needed to determine whether the trends found in this study are maintained.

Table of Contents

Overview 1
- Accommodations 3
- Examination of Validity 4

Method 5
- Participants 5
- Materials 6
  - Assessment 6
  - Accommodation 7
  - Supplemental Questions 7
- Procedure 7
- Analyses 8

Results 8

Discussion 11

Conclusion 13

References 14

Appendix A: Full Text of Directions for Standardized Administration of Form A and Form B 17

Appendix B: Full Text of Directions for Standardized Administration of Form A and Form B with the Read Aloud Accommodation 19

Overview

Over the past several years, state education boards have increasingly implemented accountability systems to ensure that schools are facilitating desired outcomes for students. In most states student performance on state-administered standardized tests is central to these accountability systems. However, for many schools, participation in these tests has included only students in the general education population. Across the nation, students in the special education system have routinely been exempted from participation in state assessments (Allington & McGill-Franzen, 1992; Langenfeld, Thurlow, & Scott, 1997; Mangino, Battaile, & Washington, 1986). Many educators believe that students with disabilities will score lower, and bring down the average performance for schools (Salvia & Ysseldyke, 1998). In fact, studies have shown that students with disabilities do score lower on standardized tests than do students without disabilities (Chin-Chance, Gronna, & Jerkins, 1996; Safer, 1980; Ysseldyke, Thurlow, Langenfeld, Nelson, Teelucksingh, & Seyfarth, 1998). With the reauthorization of the Individuals with Disabilities Education Act (IDEA) in 1997, states and districts were required to include students with disabilities in their assessments. And, exclusion has negative consequences.

First, exclusion of students with disabilities results in schools being responsible for the performance and progress of part rather than all of their students. Many state tests are high-stakes tests and are intended to be used as measures of school quality and to hold schools accountable for facilitating positive student outcomes. High-stakes tests are tests that have perceived or real consequences for students, staff, or schools (Madaus, 1988). Some states, like Texas and Maryland, have implemented policies, which include punitive consequences (e.g., state take-over of schools, firing of superintendents, etc.) for schools if their students do not perform well on statewide tests (Bushweller, 1997). States use these tests to evaluate schools and allocate resources (Langenfeld et al., 1997). According to the National Center for Education Statistics (NCES), as of the 1995-96 school year, the special education population had reached 5.6 million children, and is growing (NCES, 2000). By excluding special education students from the state tests, which are used to assess the progress and needs of students at the school level, the needs of 5.6 million students may be neglected in the resulting decisions. Exclusion means these students can be subject to an inferior education without repercussions for the school that is responsible for meeting their educational needs.

Second, in addition to being high stakes for schools and staff, these tests are also used in many cases to make high stakes decisions for individual students. Some states, such as Florida, Texas and New York require students to pass state tests in order to graduate (Langenfeld et al., 1997). For example, a student in such a state, if unable to reach a minimum level of performance on the state test, would not be eligible to receive a regular high school diploma. Instead the student might receive a certificate of completion of the Individualized Education Plan (IEP) or other similar document (Thurlow & Thompson, 1999). This certificate may not be enough to secure

the possibility for future opportunity. In 1980, Safer found that students who did not have a regular high school diploma were likely to be discriminated against in later employment. The result is that within existing state testing policies, students with disabilities want to be able to successfully participate and reap the benefits of all their many years of schooling, just as do students in the general population. However, schools are feeling pressure to exempt students from participating in these tests for fear they will earn low test scores and lower the overall district performance. In the 1997 reauthorization of IDEA, Congress took steps to resolve this conflict so that 5.6 million students are not systematically "handicapped" by state education policies.

Before students with disabilities can successfully participate in test-based accountability systems, the nature of statewide tests must be considered. Most of these tests are standardized norm-referenced tests. This means the tests are intended to be administered in a specified manner, and student performance is measured in comparison to a norm sample of students who have taken the test previously. Professional organizations argue that only with standard administration can the scores of educational and psychological tests be interpreted and student performance appropriately compared to others (American Psychological Association, 1999). Educators and advocates for students with disabilities also state that many students in the special education population are unable to take the tests in the standardized manner (Ysseldyke, 1999). For example, a student with a visual impairment may not be able to read a paper and pencil test using the standardized test booklet, and therefore would not be able to demonstrate competency through performance on the test administered in the standardized manner. There have been suggestions on how to mediate this issue. Some have suggested (1) re-norming the tests, including students with disabilities in the sample, (2) creating separate norms for students with disabilities, or (3) identifying whether specific accommodations can be made for students with disabilities without jeopardizing the validity of the test (Langenfeld et al., 1997).

The purpose of this study is to examine the third option listed above, the appropriateness of providing test accommodations to students with disabilities on norm-referenced, standardized high-stakes tests. In the reauthorization of the IDEA of 1997, the federal government considered this option as well. As mentioned above, federal lawmakers now insist that states require the participation of students with disabilities in statewide assessments. The law reads, "Children with disabilities are included in general and state wide assessment programs, with appropriate accommodations, where necessary" [IDEA, section 612(a)(17)(A)]. However, the IDEA directive leaves room for debate. It is unclear in the paragraph what is meant by an "appropriate" accommodation. It is left to educators, researchers, and test makers to work out the unanswered questions about the use of accommodations on district and state standardized tests.

Accommodations

A testing accommodation has been defined as "an alteration in the administration of an assessment" (Thurlow, Elliott, & Ysseldyke, 1998). There are many different categories of accommodations. Some authors split them into four categories (Tindal & Fuchs, 1999), others into six (Thurlow, Elliott, & Ysseldyke, 1998). The four categories included in each list are: (1) *presentation format* (e.g., reading the items to the student), (2) *response format* (e.g., marking the response in the test booklet), (3) *test setting* (e.g., in a small group), and (4) *test timing* (e.g., extended time). The purpose of a testing accommodation is to limit irrelevant sources of difficulty within the test without changing the construct measured by the test, and to facilitate access to the test by students with disabilities (Thurlow, Ysseldyke & Silverstein, 1993).

There are many issues involved with allowing students accommodations on standardized tests. These include conflict about whether the use of accommodations changes what is being measured; disagreement about whether accommodations should be used at all on standardized tests; questions of whether all accommodations used in instruction should be allowed on tests; concern about whether giving some students accommodations gives them an unfair advantage; conflict about whether accommodations should be specific to the disability; and a variety of specific issues that apply to each accommodation, without easy solutions that apply to all (Ysseldyke, 1999).

States across the country vary in how they address the issue of accommodations. Thurlow, Seyfarth, Scott and Ysseldyke (1997) indicated that 39 of the 50 states had policies on accommodations that would be allowed during statewide testing. Most allow accommodations of almost every type, but they restrict the different types of accommodations to certain portions of the tests. For example, in South Carolina, calculators are allowed on the math portion of the state's Exit Examination, but they are prohibited in the Basic Skills Assessment Program (BSAP). One state, South Dakota, does not allow any accommodations of any type on its statewide tests (Thurlow et al., 1997). One of the most controversial accommodations is reading the test aloud. Only 9 out of 39 states with accommodation policies in 1997 allow the reading aloud accommodation on all state tests with no restrictions. Sixteen states allow reading aloud on a limited number of tests and sections within tests. Three states prohibit its use entirely (Thurlow et al., 1997).

Part of the variability among states in their accommodation policies relates to one of the biggest issues with accommodations, test validity. Many accommodations are judged inappropriate by policymakers because they seem to change what the test measures. However, research has not provided clear results either way. Most studies of reading accommodations have focused on tests of mathematics. Some support the validity of reading accommodations on math tests (Fuchs, Fuchs, Eaton, Hamlet, & Karns, 2000; Tindal, Heath, Hollenbeck, Almond, & Harniss, 1998), others argue against their use (Koretz, 1997).

In terms of reading accommodations used with tests of reading and language, the literature is again thin and provides mixed results. There are strong admonitions against such accommodations, reflecting the concern that the accommodation for tests of reading decoding skills is confounded with the construct tested (McDonnell, McLaughlin & Morison, 1997). A study by Harker and Feldt (1993) reported positive effects of reading accommodations on English and content tests for high school students in the general education population. Until very recently there was no published research about the appropriateness or inappropriateness of the reading aloud accommodation on tests of reading comprehension for students with disabilities (Thurlow, Ysseldyke & Silverstein, 1993; Tindal & Fuchs, 1999). Bielinski and Ysseldyke (2000) demonstrated that reading the reading test to students with disabilities adversely affected the validity of the test. Yet, reading the math test did not adversely affect validity. Bielinski and Ysseldyke (2000) used an extant database and did not examine effects on performance of regular education students.

Examination of Validity

Many authors have approached the idea of what must be examined to determine whether test validity is spoiled. Messick (1995) introduced the idea of invalidity that is due to construct-irrelevant variance. He explained that this could be in the form of construct-irrelevant difficulty or construct-irrelevant easiness. Construct-irrelevant difficulty occurs when the task necessary for completing a test requires skills that are not related to the construct being measured, making the test especially difficult for some individuals. For example, a student with a visual impairment may not be able to read the print on a standardized test of science facts. The print size of the test would create construct-irrelevant difficulty in that the task of reading the print at that size is not related to the construct of knowledge of science facts. Accommodations, when used appropriately would serve to eliminate this type of difficulty. However, when used inappropriately, the opposite may occur. The accommodation may provide construct-irrelevant easiness. In this sense the accommodation would give the test taker outside clues in the items that may allow some individuals to choose correct answers based on information irrelevant to the construct being measured (Messick, 1995). Researchers must examine whether certain accommodations have the function of making the construct-related task easier than it would be without the accommodation.

Phillips (1994) proposed a method for sorting out appropriateness of accommodations that relates to the idea of construct-irrelevant easiness. One of the criteria included in the method she proposes states that scores from accommodated tests should not be considered valid if the accommodation has similar effects on scores of students with and without disabilities. An accommodation that results in elevated scores for both groups of students may create construct-irrelevant easiness. If the effect of the accommodation is the same for both groups, it is implied

that the accommodation is providing clues that are unrelated to the construct irrelevant difficulty, which the accommodation is attempting to eliminate. Although this is only one factor to consider when determining validity, it is a stepping-stone for reaching a final conclusion about the appropriateness of specific accommodations.

Methods for evaluating the effect that accommodations have on level of construct related difficulty and test validity have been discussed. Tindal (1998) proposed three different methods of examining validity in conjunction with test accommodations: descriptive, comparative and experimental. The experimental model was chosen for this study. Within the experimental model, equal size groups of both students with disabilities and without disabilities are given the same test. They are also given the test under two conditions, with accommodations and without. This design allows for the control of internal sources of variance. With this method it is possible to conduct an experimental examination of the effects of the accommodation on both groups of students and compare the effects with one another, to determine whether the effects are similar or different. As stated above, this is a beginning point in a series of evaluations necessary to conclude that reading accommodations do not violate the validity of tests of reading comprehension.

In this study, the following questions were addressed:

1. To what extent does reading the reading comprehension test affect the scores of students with disabilities?

2. To what extent does reading the reading comprehension test affect the scores of students in the general population?

3. Are there differences in the effects on the two populations?

Method

Participants

This study was conducted in a suburban school district outside of a major midwestern city. Seventy students were recruited from grades 3 though 5, in five elementary schools. Thirty-two agreed to participate. Seventeen students were in the general education population and 15 students were part of the special education population. Efforts were made to keep the groups as comparable as possible in terms of demographic characteristics. However, due to the limited number of students willing to participate, the special education group was comprised mostly of males. A description of demographic characteristics is provided in Table 1.

Table 1. Demographic Characteristics of Participants in the General Education and the Special Education Groups

	General Education Group	Special Education Group
Gender		
Male	8	12
Female	9	3
Grade		
3	4	3
4	2	4
5	9	7
Ethnicity		
African American	3	2
European American	11	11
Other	3	2

Students in the regular education sample represented varied reading skill levels from low to high, as verified by teacher accounts of in-class student performance. Students in the special education group all were currently receiving special education services in reading. Nine of the students were receiving services under the learning disability (LD) category. Four students were receiving services under the category of emotional/behavior disorder (E/BD), and one student was receiving services under the speech and language category. All students were identified by the schools as having reading difficulties and were receiving reading services through the special education system.

Materials

Assessment

The California Achievement Tests, Fifth Edition (CAT/5) is a set of tests designed to measure achievement of basic skills in the areas of reading (vocabulary, comprehension), math (computation, concepts and application), science, language (mechanics, expression), spelling, study skills, and social studies. From these tests it is possible to obtain mastery scores for each specific objective. The tests are configured two different ways: a Survey Battery and a Complete Battery. For this study, the Survey Battery was chosen, since it includes all the skill areas, but is much shorter than the Complete Battery. For the purpose of this study, students were given the Comprehension test only. [California Achievement Tests, 5th Edition by CTB/McGraw-Hill, a division of The McGraw-Hill Companies, 20 Ryan Ranch Road, Monterey, CA 93940. Copyright © 1992 by CTB/McGraw-Hill. All rights reserved.]

The Comprehension test within the survey battery contains 20 items. On this test students are required to read passages and answer questions about the passage, including pulling out details,

interpreting events, and analyzing characters. They are also required to analyze writing in terms of fact and opinion and extended meaning. There is a twenty-minute time limit for this test.

The CAT/5 Survey Battery is available in 13 different levels. Level 14 was used because it is recommended by the test publisher for students in the middle of third grade through the beginning of fifth grade. The Survey is also available in two equivalent forms, Form A and Form B. Students took both forms of the Comprehension test. One form was administered in the standardized format prescribed by the test publisher. The other form was administered with the test accommodation. A counterbalanced design was used.

Accommodation

To maintain consistency between testing sessions, the read aloud accommodation was provided using a standard audiocassette player. Prior to the testing sessions, each form of the Comprehension test was recorded on a cassette tape. One male and one female adult reader alternated between each passage and set of items for each test form. Each passage was read once, at a speed of approximately 120 words per minute, followed by the corresponding items. For each item, the question was read once, followed by the response choices, and then read a second time. After the question was read the second time, the reader paused 10 seconds before moving on to the next item.

Supplemental Questions

Two open-ended questions were asked of the students at the end of the testing session to get an idea of student perception of and comfort level with the read aloud test accommodation. They were asked: (1) Which way of taking the test did you like better, with the tape or without the tape? (2) Why did you like that way better?

Procedure

Testing sessions were conducted in the spring and fall of 1999. Students were tested in groups of one to five students. Students in the general education group and students in the special education group were tested separately. Each student participated in one testing session. During each testing session, both Form A and Form B were administered. All students took Form A first and Form B second. However, to minimize error due to practice, the order of the accommodation was counterbalanced among the students. In other words, half of the students took Form A with the tape first and Form B without the tape second. The other half took Form A without the tape first and Form B with the tape second.

Each testing session began with an introduction to the testing session, an overview of the procedure and an opportunity for questions. The students then took Form A of the test, followed by a short break. Following the break, students took Form B of the test, followed by answering the supplemental questions. Students recorded all their answers directly in the test booklet. Having students mark answers in the test book instead of an answer sheet is not standard procedure for the CAT/5. However, studies have shown that whether students mark answers in the test book or on the answer sheet does not affect the resulting scores (Rogers, 1983; Tindal et al., 1998).

For the administration of both Forms A and B without the tape, the procedure outlined in the examiner's manual was followed. However, in the directions, some of the wording was changed to reflect the request that students mark answers directly in the test book. For the full text of the directions for standard administration, see Appendix A. Students were given a twenty minute time limit for the without the tape condition. For the administration of both Form A and B with the tape, the general procedure outlined in the examiner's manual was followed. However, substantial changes were made to the direction text. For full text of directions used for the accommodation condition see Appendix B. Instead of a time limit on the test, students were expected to follow along with the tape and not go ahead. A few times during the sessions it was necessary for the examiner to stop the cassette tape to allow a student to complete a response, but this was generally not needed for most of the students.

Analyses

Each student's score for performance on each test was converted from the raw score to a scaled score and percentile rank based on the national norms provided by the testing company. Multiple analyses of the difference between means as well as descriptives were used to evaluate effects. Analyses of the supplemental questions included qualitative description of trends and degree of congruency between student perception of effects of the accommodation and actual scores.

Results

Complete data sets were collected for all 17 students in the general education group and 14 students in the special education group. The scores for one of the students in the special education group were not used due to a substantial interruption in the testing session. Student test performance is reported in both scaled scores and percentile ranks. All statistical analyses were conducted on scaled scores. Percentile ranks are also reported and discussed for ease in understanding effects.

In the first analysis, the difference between mean scores for the special education group was analyzed. The performance of students with disabilities on the administration without the accommodation yielded a mean scaled score of 661.4. This group's performance on the test, when administered with the accommodation, yielded a mean scaled score of 691.6. The difference between these means approaches statistical significance ($t_{(13)}$=2.09, p=.06). Although not statistically significant, the difference in means yields a moderate effect size (0.56).

In terms of percentile ranks, the mean percentile rank for students with disabilities on the standard administration was 28.7. The mean percentile rank when the test was administered with the accommodation was 48.6. The average gain for the group in percentile ranks was 19.9. The range of gains in percentile ranks was –17 through 64.

To address the second question on the effect of the accommodation on scores for students in the general education population, an analysis of difference between scaled score means was computed. The mean scaled score for the students without disabilities when the test was administered with the standardized procedure was 744.6. The mean when the accommodation was added to the procedure was 749.8. The difference in the two means is not statistically significant ($t_{(16)}$=0.68, p=.69). In terms of percentile ranks, students in the general population group earned an average percentile rank of 67.9 when the test was presented without the accommodation. They earned, on average, a percentile rank of 72.5 when the test was administered with accommodation. The range of gains in percentile ranks was from –49 to 65.

The third research question addressed in this study was whether there was a difference in the effects of the accommodation on students' performance based on student status. A repeated-measures analysis of variance was computed. The interaction between the two was not significant. The difference between group means is illustrated in Figure 1.

In addition to statistical analyses, a qualitative analysis of students' responses to the supplemental questions was also was conducted. Answers to the supplemental questions were obtained from all 17 of the students in the general education group and 12 of the students in the special education group. More students in the special education group indicated that they preferred the test administration with the accommodation (with = 75%, without = 25%). In contrast, the students in the general education group preferred the test administration without the accommodation (without = 76%, with = 23%). The majority (66%) of the students in the special education group who preferred the accommodation procedure indicated that they preferred this way because it reduced the difficulty of the task. For students in the general education group, who preferred the administration without the accommodation, most of them responded either that they liked that they could go at their own pace or specifically because it was "faster." For a detailed list of student responses see Table 2.

Figure 1. Difference in Performance by Administration Procedure for Students in the General Education and Special Education Groups

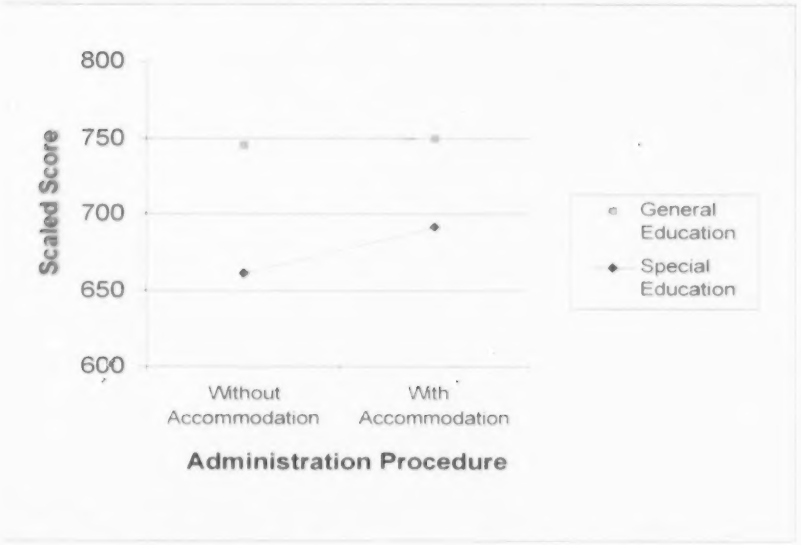


Table 2. Summary of Student Responses to Supplemental Questions

Question 1: Which way of taking the test did you like better? With the tape or without?

General Education Group*
With the tape – 23%
Without the tape – 76%

Special Education Group*
With the tape – 75%
Without the tape – 25%

Question 2: Why did you like that way better?

General Education Group**
Without the tape:
"I could go at my own pace." – 6
"It was faster." – 6
"I didn't like the tape." – 1

With the tape:
"I didn't have to read." – 2
"I didn't have to pay attention." – 1
"It was easier." - 1

Special Education Group**
Without the tape:
"It was faster." – 2
"The stories were better" -- 1

With the tape:
"It was easier." – 6
"It was faster." – 2
"I didn't have to read." – 1
"The stories were better." – 1

* Percentages indicate the percent of students in the group who gave this response.
** Numbers indicate the total number of students in the group who gave this response.

The final analysis was conducted on the amount of agreement between the students' preference and their actual performance in terms of scaled scores. For students in the special education group, 83% of the students actually performed better on the administration procedure that they preferred. However, only 41% of the students in the general education group performed better on the test when it was administered the way they preferred.

Discussion

This study focused on the effect of providing a reading accommodation on the validity of reading comprehension tests. Students in the general education population and students in the special education population were both given two equivalent forms of a reading comprehension test. Each student took one form of the test under standard procedures, and the other form of the test with the reading accommodation. The difference in the performance of students under each testing procedure was compared for both groups. Information was also collected from students in each group about their perception of the accommodation.

The students' scores in this study did not reveal a statistically significant interaction between the accommodation and student group. According to Phillips (1994), one of the criteria that must be met for an accommodation to be considered appropriate is that it must not have the same effect on the performance of students with disabilities and without. If the accommodation increases scores for both groups of students, it is possible that the change introduces, what Messick (1995) defines as construct-irrelevant easiness. Messick argues that introducing construct-irrelevant easiness to a testing procedure invalidates the resulting test scores. Even though no statistical interaction was found, the absence of a statistical interaction should be interpreted with caution.

In order to have a significant interaction, at least one of the variables must reveal a significant simple effect. The performance of students in each group was different, but the differences were not statistically significant in this study. The effect of the accommodation on the general education group was small and far from significant ($d=0.10$, $p=0.69$). However, the effect on the special education group was much more substantial and close to statistical significance ($d=0.56$, $p=0.06$). The failure of this study to produce a statistically significant effect for the accommodation for students with disabilities may be a function of the small sample size of the groups. With a reasonably larger sample size for both groups, the difference in the effect of the accommodation for students by group may reach significance.

While establishing statistical significance is critical for reaching conclusions on the appropriateness of accommodations, clinically significant differences are important to consider as well. A clinically significant difference can be conceptualized as a difference that is large enough to impact decisions made about an individual or group. The difference in performance

for the students in the special education group may not be statistically significant, but may be clinically significant. The average difference in scaled scores for students with disabilities between the accommodated version and the non-accommodated version was 30.2, when converted to percentile ranks, this difference translates to an average jump of 19.9 percentile ranks. For example, based on the trend in this study, if a student in the special education population earns a percentile rank of 15 when the test is given without the accommodation, that student may have earned a percentile rank of 35 if the test were given with accommodations. While not statistically significant in this study, an average difference this size can have a substantial impact on educational decisions and on students' lives.

In contrast to the large effect the accommodation had for students with disabilities, the effect of the accommodation for students in the general education population was quite small. The average increase in scaled score units was only 5.2. In terms of percentile ranks, the increase was only 4.6. An average difference between the 63rd percentile and the 67th percentile probably would have insignificant educational impact for individual students as well as for schools as a whole. In this study a significant interaction effect of the accommodation and student group was not demonstrated. However, the trend demonstrated by the results indicates a possible clinical effect of the accommodation for students with disabilities that is not present for students without disabilities. As mentioned above, further investigation of this accommodation could shed more light on the presence of this possible differential effect.

In addition to analysis of the effect of the accommodation on student's scores, student perception of the accommodation was qualitatively analyzed. There was a difference between the groups in terms of which procedure they preferred. The majority of students in the general education group preferred taking the test without the accommodation. They complained that the accommodation slowed down the test too much and did not allow them to proceed at their own pace. This suggests that students in the general education population may not want this accommodation and actually find it aggravating. On the other hand students in the special education population welcomed the accommodation. Many shared that reading is difficult for them and the accommodation helped them understand the questions, making the test less difficult. These students know they are working on developing their reading skills and believe the accommodation was helpful to them.

In the final analysis, student's perceptions were compared to their performance on the test with and without the accommodation. In general, students with disabilities were more accurate in that they preferred the testing procedure with which they performed better. The reason for this greater congruence may be explained by the uniformity of the accommodation effects for students in this group. Of the 12 students who responded to the questions, 11 of them performed better with the accommodation and only one performed worse. Given that 9 of them stated that they preferred the accommodation procedure, the agreement between their performance and preference

is not surprising. The performance of the students in the general education population was more variable. Nine students' scores went up with the accommodation, seven students' scores when down and one stayed exactly the same. However, most of these students preferred the test procedure without the accommodation. Thus, there was less congruence between performance and preference for students in the general education population.

Conclusion

The purpose of this study was to investigate the effects of a reading accommodation on the performance of students with and without disabilities on a reading comprehension test. A significant interaction of test scores, indicating differential effects of the accommodation by group was not demonstrated. However, the effect sizes for each group were substantially different. The lack of a statistically significant interaction may simply be a function of the small sample size of this study. Further investigation using larger groups is needed. Moreover, student perception of the accommodation was different based on group membership. Students in the special education group embraced the accommodation, while those in the general education group rejected it.

As mentioned earlier, establishing a differential effect of accommodations for students with and without disabilities is only a beginning step in determining the appropriateness of test accommodations. Particularly, with such controversial accommodations as reading a reading test, a more in depth examination of test validity is required before conclusions can be reached. The information in this study can provide insight for future investigations of reading accommodations and put researchers one step closer to weeding through the complexity of identifying ways in which the growth and progress of all students can be measured in America's schools.

References

Allington, R., & McGill-Franzen, A. (1992). Improving school effectiveness or inflating high-stakes test performance? *ERS Spectrum, 10*(2), 3-12.

American Psychological Association (1999). *Standards for educational and psychological testing*. Washington, D. C.: American Educational Research Association

Bielinski, J., & Ysseldyke, J. (2000). *The impact of testing accommodations on test score reliability and validity*. Paper presented at the 30th Annual National Conference on Large-Scale Assessment, Snowbird, UT.

Bushweller, K. (1997). Teaching to the test. *American School Board Journal*, 184 (9), 20-25.

Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlet, C., & Karns, K. (2000). Supplemental teacher judgments about test accommodations with objective data sources. *School Psychology Review*, *29* (1), 65-85.

Harker, J. K., & Feldt, L. S. (1993). A comparison of achievement test performance of nondisabled students under silent reading and reading plus listening modes of administration. *Applied Measurement, 6*, 307-320.

Koretz, D. (1997). *Assessment of students with disabilities in Kentucky* (CSE Technical Report No. 431). Los Angeles, CA: Center for Research on Standards and Student Testing.

Langenfeld, K. L., Thurlow, M. L., Scott, D. L. (1997). *High-stakes testing for students: Unanswered questions and implications for students with disabilities* (Synthesis Report 26). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Madaus, G. (1988). *The influence of testing on the curriculum*. In L. Tanner (ed.), Critical issues in curriculum: 87th Yearbook of the NSSE Part 1. Chicago, IL: University of Chicago Press (ERIC Document Reproduction Service No. 263 183).

Mangino, E., Battaile, R., & Washington, W. (1986). *Minimum competency for graduation* (Publication Number 84.59). Austin, Texas: Austin Independent School District. (ERIC Document Reproduction Service No. 263 183)

McDonnell, L. M., McLaughlin, M. J., Morison, P. (1997). *Educating one and all: Students with disabilities and standards-based reform*. Washington, D. C.: National Academy Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

National Center for Education Statistics (2000). *Fast facts*. Web site, http://www.nces.gov/fastfacts/.

Phillips, S. E. (1994). High stakes in testing accommodations: Validity versus disabled rights. *Applied Measurement in Education, 7*(2), 93 -120.

Rogers, W. T. (1983). Use of Separate answer sheets with hearing impaired and deaf school age students. *B. C. Journal of Special Education, 7*(1), 63-72.

Safer, N.D. (1980). Implications of minimum competency standards and testing for handicapped students. *Exceptional Children, 46*, 288-290.

Salvia, J., & Ysseldyke, J. E. (1998). *Assessment*. (Seventh Edition) Boston: Houghton Mifflin Company.

Thurlow, M. L., Elliott, J., & Ysseldyke, J. E. (1998). *Testing students with disabilities: Practical strategies for complying with district and state requirements*. Thousand Oaks, CA: Corwin Press.

Thurlow, M. L., Seyfarth, A., Scott, D., & Ysseldyke, J. (1997). *State assessment policies on participation and accommodations for students with disabilities: 1997 update* (Synthesis Report 29). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M., & Thompson, S. (1999). *Diploma options and graduation policies for students with disabilities* (Policy Directions No. 10). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M., Ysseldyke, J., & Silverstein, B. (1993). *Testing accommodations for students with disabilities: A review of the literature* (Synthesis Report 4). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Tindal, G. (1998). *Models for understanding task comparability in accommodated testing*. Eugene, OR: University of Oregon, SCASS-ASES study group.

Tindal, G., & Fuchs, L. (1999). *A summary of research on test changes: An empirical basis for defining accommodations*. University of Kentucky: Mid-South Regional Resource Center Interdisciplinary Human Development Institute.

Tindal, G., Heath, B., Hollenbeck, K., Almond, P., & Harniss, M. (1998). Accommodating students with disabilities on large-scale tests: An empirical study of student response and test administration demands. *Exceptional Children, 64* (4), 439-450.

Ysseldyke, J. E. (1999). *Assessment and accountability systems: Current status and thoughts about directions.* (University Presentation). Minneapolis: University of Minnesota, National Center on Educational Outcomes

Ysseldyke, J. E., Thurlow, M. L., Langenfeld, K., & Nelson, J. R. Teelucksingh, E., & Seyfarth, A. (1998). *Education results for students with disabilities: What do the data tell us?* (Technical Report 23). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Appendix A

Full Text of Directions for Standardized Administration of Form A and Form B

Appendix A

Full text of directions for standardized administration of Form A and Form B.

Open your book to test 2, Comprehension, beginning on page 7.

This test is about understanding what you read. When you mark an answer, *circle the letter** that goes with the answer you choose. If you want to change an answer, erase the *circle* you made and make a new *circle*. We will do one sample item. When you start working on the test, be sure to read the directions in your test book for each set of items. Find the sample item, Sample A.

For Sample A, read the passage. Then read the questions below the passage. Find the answer. *Circle the letter* that goes with the answer you choose for Sample A and then stop.

For Sample A, you should have *circled* D. D goes with the answer "oranges." From the passage, you know that Barry sold oranges in the winter. If you did not *circle* D, erase the *circle* you made and *circle* the correct answer.

Now you are going to do some more items by yourself. Remember to read all the directions carefully. When you see the words go on at the bottom of the page, go right on to the next page. When you come to the stop sign, you have finished this test. When you finish you may check your answers. Then sit quietly until the other students have finished. You will have 20 minutes to do this test. Are there any questions?

* Italic print indicates a word change from the published version.

Appendix B

Full Text of Directions for Standardized Administration of Form A and Form B with the Read Aloud Accommodation

Appendix B

Full text of direction for administration of Form A and Form B with the read aloud accommodation.

Open your book to test 2, Comprehension, beginning on page 7.

This test is about understanding what you read. When you mark an answer, *circle the letter** that goes with the answer you choose. If you want to change an answer, erase the *circle* you made and make a new *circle*. We will do one sample item. When you start working on the test, be sure to read the directions in your test book for each set of items. Find the sample item, Sample A.

For Sample A, you will hear the reader read all the words on the page. Please read along to yourself then find the answer.** *Circle the letter* that goes with the answer you choose for Sample A and then stop.

For Sample A, you should have *circled* D. D goes with the answer "oranges." From the passage, you know that Barry sold oranges in the winter. If you did not *circle* D, erase the *circle* you made and *circle* the correct answer.

Now you are going to do some more items by yourself. Read along to yourself as the reader reads the words on the page. I will stop the tape between items to allow you to circle your answer. When you finish marking your answer, please wait until the next item is read, or wait until directions are given before moving on. Are there any questions?

* Italic print indicates a word change from the published version.

** Underlined print indicates changes specific to the accommodation condition.

The College of Education
& Human Development

UNIVERSITY OF MINNESOTA